EduCloud - Implantação #530

Implantação # 519 (Resolvida): US: Investigação Preliminar do Hadoop

Analise de viabilidade da tecnologia

09 Apr 2012 09:19 - Marco Gomes

Status:	Resolvida	**Start date:**	09 Apr 2012
Priority:	Normal	**Due date:**	10 Apr 2012
Assignee:	Marco Gomes	**% Done:**	100%
Category:		**Estimated time:**	2.00 hours
Target version:	Ciclo 2	**Spent time:**	2.50 hours

Description

A tecnologia faz sentido para o projeto ? E quais vantagens e desvantagens da tecnologia ?

History

#1 - 09 Apr 2012 09:37 - Marco Gomes

- Start date changed from 10 Apr 2012 to 09 Apr 2012

- Estimated time changed from 4.00 to 2.00

#2 - 18 Apr 2012 11:01 - Marco Gomes

- Status changed from Nova to Execução

- % Done changed from 0 to 40

Dado que não está claro que o projeto EduCloud irá trabalhar com grandes volumes de dados, é provável que o hadoop não seja a solução mais indicada para o projeto, pois todo o trabalho existente no hadoop envolve a movimentação e disponibilidade de grandes volumes de dados.

A principio, as vantagens da tecnologia são disponibilizar o dado com redudância, tentar manter o dado próximo ao local onde ele será usado para evitar uso de rede (sobrecarga nos switches centrais) e realizar o correlacionamento de dados semelhantes.

A desvantagem do hadoop é ser um filesystem baseado em write-once, ou seja, não existe append de arquivo, não é possível modificações de arquivos, e provável que ele não tenha performance ideal com baixo volume de dados.

#3 - 20 Apr 2012 13:03 - Marco Gomes

Outra desvantagem é o map reduce. O map reduce se trata de uma linguagem para criar jobs de análise de dados. É necessário um conhecimento prévio do que se deseja analisar, pois o hadoop apenas fornece as ferramentas para análise, não as faz de forma automática. Portanto deve se levar em consideração o tempo de desenvolvimento dos jobs de map reduce.

#4 - 20 Apr 2012 13:05 - Marco Gomes

- % Done changed from 40 to 100

#5 - 20 Apr 2012 13:06 - Marco Gomes

- Status changed from Execução to Resolvida